INFECTION AND IMMUNITY, Nov. 1989, p. 3281–3291
0019-9567/89/113281-11$02.00/0

Vol. 57, No. 11

DNA Sequence and Expression of the 36-Kilodalton Outer Membrane Protein Gene of *Brucella abortus*

THOMAS A. FICHT,[1,2]* SCOTT W. BEARDEN,[1] BLAIR A. SOWA,[3] AND L. GARRY ADAMS[3]

Departments of Veterinary Microbiology and Parasitology,[1] Genetics,[2] and Veterinary Pathology,[3] Texas A&M University and Texas Agricultural Experiment Station, College Station, Texas 77843

Received 23 March 1989/Accepted 17 July 1989

The cloning of the gene(s) encoding a 36-kilodalton (kDa) cell envelope protein of *Brucella abortus* has been previously described (T. A. Ficht, S. W. Bearden, B. A. Sowa, and L. G. Adams, Infect. Immun. 56:2036–2046, 1988). In an attempt to define the nature of the previously described duplication at this locus we have sequenced 3,500 base pairs of genomic DNA encompassing this region. The duplication represented two similar open reading frames which shared more than 85% homology at the nucleotide level but differed primarily because of the absence of 108 nucleotides from one of the two gene copies. These two genes were read from opposite strands and potentially encoded proteins which are 96% homologous. The predicted gene products were identical over the first 100 amino acids, including 22-amino-acid-long signal sequences. The amino acid composition of the predicted proteins was similar to that obtained for the *Brucella* porin isolated by Verstreate et al. (D. R. Verstreate, M. T. Creasy, N. T. Caveney, C. L. Baldwin, M. W. Blab, and A. J. Winter, Infect. Immun. 35:979–989, 1982) and presumably represented two copies of the porin gene, tentatively identified as *omp* 2a (silent) and *omp* 2b (expressed). The homology between the two genes extended to and included Shine-Dalgarno sequences 7 base pairs upstream from the ATG start codons. Homology at the 3′ ends extended only as far as the termination codon, but both genes had putative rho-independent transcription termination sites. Localization of the promoters proved more difficult, since the canonical procaryotic sequences could not be identified in the region upstream of either gene. Promoter activity was demonstrated by ligation to a promoterless *lacZ* gene in pMC1871. However, only one active promoter could be identified by using this system. A 36-kDa protein was synthesized in *E. coli* with the promoter in the native orientation and was identical in size to the protein produced in laboratory-grown *B. abortus*. When the promoter-containing fragment was inverted, a 33-kDa protein was expressed. These results were consistent with the predicted sizes based on the nucleotide sequences of the open reading frames in *omp* 2b and *omp* 2a. Whether this locus contains one active and one silent or cryptic porin gene, or two active *Brucella* porin genes expressed under different environmental conditions, is discussed.

Little is known about the molecular mechanisms used by the facultative intracellular bacterium *Brucella abortus* to penetrate and resist destruction within the host macrophage. Studies have focused primarily on the lipopolysaccharide component of the outer membrane (12, 27). More recently, low-molecular-weight nucleotides thought to inhibit degranulation have been described (2, 3). Alterations in the structural composition or metabolism of smooth organisms which may accompany changes in environmental conditions have not been investigated since the work of Frost et al. (8). At that time these investigators suggested that the increased virulence of in vivo-grown organisms in a guinea pig model was due to a cell wall component distinct from the major antigenic lipopolysaccharide. In addition, *B. abortus* has been shown to preferentially utilize the four-carbon aliphatic alcohol erythritol, which is present at high concentrations in infected bovine tissues (19). Our goal was to determine any changes which potentially occur in *B. abortus* as a result of changing environmental conditions during the course of infection and which play a role in intracellular survival and virulence.

We have previously reported the cloning of the gene(s) encoding a 36-kilodalton (kDa) protein identified in the cell envelope fraction of *B. abortus* (7). Additional evidence indicated its association with the outer membrane, thus circumstantially suggesting the cloning of the *B. abortus* porin genes. It was shown that sequences encoding an antigenic portion of the 36-kDa protein are repeated within a 3.5-kilobase-pair (kbp) stretch of the *B. abortus* genome. An oligonucleotide probe specific for the amino-terminal end of the mature gene product also hybridizes to multiple fragments in restriction digests of the cloned *Brucella* locus. It was unclear, however, whether the duplicated sequences are contained within a single gene or represent two gene copies.

This study reports the sequence of the 3.5-kbp stretch of *B. abortus* genomic DNA containing the 36-kDa protein gene locus and demonstrates that a second open reading frame, potentially encoding a related protein, is present. Comparison of the amino acid sequence of the predicted proteins with that obtained by Verstreate et al. for purified *Brucella* porins is consistent with the identification of the porin gene locus *omp* 2 of *B. abortus* (35). The transcriptional and translational activities of both genes in *Escherichia coli* are described. It is unclear, however, on the basis of their expression in *E. coli*, whether both copies represent functional *Brucella* genes. The genetic arrangement of these *Brucella* genes in relation to other divergent promoters is discussed, as are the implications regarding the potential for expression of porins of altered selectivities under different environmental conditions.

MATERIALS AND METHODS

Bacterial strains and cultivation. *B. abortus* smooth strains 19 and 2308 were obtained from Billy Deyoe at the National

* Corresponding author.

Animal Disease Center in Ames, Iowa. *Brucella* species designations for all strains were confirmed by standard biotype analysis (1). *E. coli* JM105 [*thi rpsL endA sbcB15 hsdR4* Δ(*lac-pro* AB)/F′ *traD36 proAB lacI*^{q} ZΔM15], DH5αF′ [λ$^-$*recA*1 *endA*1 *hsdR17* (r$^-_\kappa$ m$^+_\kappa$) Δ(*lacZYA-argF*)*U169* φ80d *lac*ZΔM15 *supE44 thi-1 gyrA96 relA*1], and MV1190 [Δ(*lac-pro*) *thi supE*Δ(*sr-1-recA*)::Tn*10*(Tet^r) F′ *traD proAB lac*I^q ZΔM15] were obtained from Pharmacia Pharmaceuticals, Bethesda Research Laboratories, Inc. and Bio-Rad Laboratories, respectively. *E. coli* MC4100 [*F*$^-$ *araD139* Δ(*argF-lac*)*U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR*] was obtained from M. J. Casadaban through Ry Young in the Department of Biochemistry and Biophysics, Texas A&M University. *E. coli* SE5000 [MC4100 (*recA*)] was obtained from Karin Ihler in the Department of Medical Microbiology, Texas A&M University. *E. coli* ECB611 [MC4100(*lamB ompF*::Tn*5 ompC*::Tn*10*)] was obtained from Spencer Benson in the Department of Microbiology at the University of Maryland. *E. coli* LE392 (F$^-$ *supF supE hsdR galK trpR metB lacY tonA*) and P2 392 [LE392(P2)] were obtained from Stratagene, Inc. Bacteriophage M13mp18 and mp19 replicative-form DNA were obtained from Bethesda Research Laboratories, Inc. *B. abortus* and *E. coli* were grown as previously described (7).

Subcloning *B. abortus* genomic DNA. Recombinant libraries were prepared and recombinants expressing Omp 2 antigens were selected as previously described (7). Larger genomic recombinants in λ2001 were selected by using the λgt11 inserts subcloned into M13 (7). Southern blot analysis was used to identify DNA fragments which hybridized to both the M13 and oligonucleotide probes (7). A single *Bam*HI restriction fragment of 6.5 kbp originating from *B. abortus* S2308 was subcloned into pBR322 at the unique *Bam*HI site, and the resulting plasmid was designated pAGF101. The corresponding fragment from *B. abortus* 19 (3.5 kbp) was excised from λ2001 by digestion with *Bam*HI and *Xho*I and subcloned into pBR322 at the *Bam*HI and *Sal*I sites and designated pAGF201. The difference in the sizes of the fragments derived from S19 and S2308 is a cloning artifact caused by partial digestion of the *Brucella* genomic DNA with *Sau*3AI for the purpose of library construction. The 6.5-kbp *Bam*HI DNA fragment has been found to be conserved in the genomic DNA of both the vaccine (S19) and virulent (S2308) strains (7). For simplicity, only the 3.5-kbp region common to both the S19 recombinant (pAGF201) and S2308 recombinant (pAGF101) will be described in detail here. *Pst*I fragments were subcloned from pAGF101 or pAGF201 into pUC9 (18, 36) or pMC1871 (4), which were selected because of the correct alignment of reading frames between the *omp* 2 translation initiation sites and *lacZ*. The location of insertion and the fragment orientation were performed by restriction enzyme analysis. In pMC1871, *Apa*I-*Eco*RI digests were used to identify the *Pst*I site at which the *Brucella* DNA fragment was inserted. The orientation was determined by digestion with *Hin*cII or *Acc*I, which cut asymmetrically in the 350-bp fragment and the 550-bp fragment, respectively.

Restriction mapping of the *omp* 2 locus. The *B. abortus* insert was mapped by single, double, and triple restriction enzyme digestion and Smith and Birnstiel mapping (32). In the latter case, the appropriate fragments were end labeled by using T4 DNA polymerase as described by Maniatis et al. (15) and digested a second time with the appropriate enzyme, and the end-labeled fragments were purified by electroelution.

Subcloning of DNA fragments into M13 mp18. Recombinant plasmids pAGF101 and pAGF201 containing *B. abortus* DNA inserts were digested with restriction endonuclease *Pst*I, *Rsa*I, *Pst*I-*Eco*RI, *Pst*I-*Bam*HI, or *Rsa*I-*Hin*cII, and the fragments were purified by electroelution and ligated into M13 mp18, all as previously described (7). DNA sequences were determined for both strands by the dideoxy method with 7-deazaGTP (Boehringer Mannheim Biochemicals) in place of dGTP to eliminate compression problems due to GC-rich regions. Hybridization probes specific for the *B. abortus* inserts were prepared by using the single-stranded DNAs as templates and were labeled as described by Hu and Messing (11).

Preparation of *B. abortus* genomic DNAs for Southern blot analysis. Genomic DNAs were extracted as described previously (7) and digested overnight with restriction enzymes (5 to 10 U/μg of DNA) in 100-μl reaction volumes at the appropriate temperature, in buffer supplied by the manufacturer. Transfer to nylon membranes (Zeta-Bind; CUNO Inc.), prehybridization, hybridization, and washing were performed as previously described (7).

Protein synthesis and identification by Western blot (immunoblot) analysis. Recombinant plasmids pAGF101 and pAGF201 were digested with *Asp* 718 and were subsequently religated. *Asp* 718 (Boehringer Mannheim) is an isoschizimer of *Kpn*I which generates 5′-protruding ends. *Kpn*I sites are present within the 5′ end of both structural genes, *omp* 2a and *omp* 2b. Recombinants were selected and characterized by restriction enzyme analysis. Plasmids containing the *Kpn*I fragment in the native orientation (pAGF101 and pAGF201), inverted orientation (pAGF111 and pAGF211), and plasmids lacking the *Kpn*I fragment (pAGF11 and pAGF21) were transformed into competent *E. coli* ECB611. The production of *omp* 2 gene products was examined in mid-log-phase cultures grown to an optical density at 550 nm of 0.5, and harvested by centrifugation at 4,000 × *g*. The cells were lysed by sonication and clarified by centrifugation at 12,000 × *g*. The cell envelope fraction was obtained by centrifugation at 100,000 × *g* (14), suspended in Laemmli sample buffer, and boiled for 5 min just prior to electrophoresis. Electrophoresis was performed on an 8 to 14% (wt/vol) polyacrylamide gradient gel, and either the gel was stained with Coomassie brilliant blue (13) or the proteins were transferred to nitrocellulose for Western blot analysis (33). Detection of the *omp* 2 gene products on Western blots was performed by using rabbit antisera as primary and alkaline phosphatase-conjugated goat anti-rabbit (immunoglobulin G) as secondary antibody, as described previously (7).

Maxicell synthesis of *B. abortus omp* 2 gene products. Growth and preparation of the maxicells was performed essentially as described by Sancar et al. (29) with the following modifications. Plasmids were transformed into competent *E. coli* SE5000 and grown in M9 minimal media (15). The cells were pulse-labeled for 1 h with 25 μCi each of [^{3}H]glycine (specific activity of 20 Ci/mmol) and [^{3}H]leucine (specific activity of 130 Ci/mmol), pelleted by centrifugation at 10,000 × *g* for 5 min, and suspended in 100 μl of Laemmli sample buffer. Portions (25 to 50 μl) were loaded onto a sodium dodecyl sulfate-polyacrylamide gel and electrophoresed as described above. The gel was soaked in En^{3}Hance (Dupont, NEN Research Products) and dried at 60°C in vacuo. The dried gel was exposed to X-ray film (XAR-5; Eastman Kodak Co.) for 24 to 48 h.

RESULTS

Restriction mapping. Genomic recombinants in λ2001 containing *Brucella* DNA inserts of approximately 20 kbp were

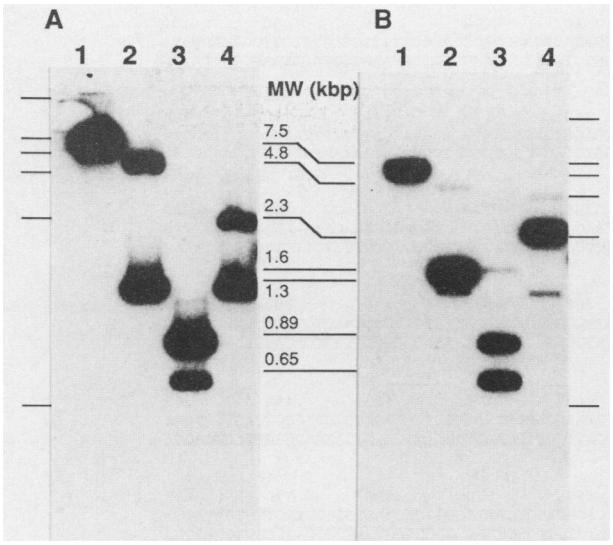


FIG. 1. Southern blots of *B. abortus* recombinant plasmid containing the *omp* 2 locus. Recombinant plasmid pAGF201 DNA was digested with *Bam*HI (lanes 1), *Eco*RI (lanes 2), *Pst*I (lanes 3), and *Cla*I (lanes 4). (A) Hybridization with λgt11 probe; (B) hybridization with oligonucleotide probe. For description of plasmids and probes, see text. Horizontal lines indicate the relative migration of *Hin*dIII-cut λ DNA fragments of 23, 9.4, 6.7, 4.4, 2.3, and 0.5 kbp, from top to bottom.

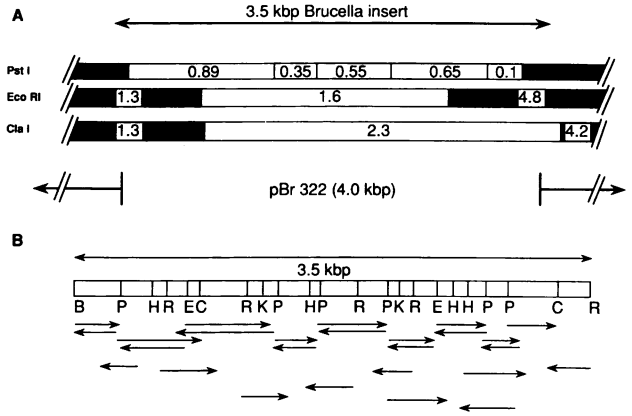


FIG. 2. Restriction map and sequencing strategy of *B. abortus* *omp* 2 locus. (A) *Pst*I, *Eco*RI, and *Cla*I restriction maps of the *omp* 2 locus in pAGF201 which contains a 3.5-kbp *B. abortus* genomic DNA insert. The stippled boxes represent joint fragments containing both *Brucella* and pBR322 DNAs produced by digestion with the restriction enzymes. The sizes of the fragments are shown in kilobase pairs. (B) Composite restriction map. B, *Bam*HI; C, *Cla*I; E, *Eco*RI; H, *Hin*cII; K, *Kpn*I; P, *Pst*I; R, *Rsa*I. Arrows indicate the sequencing strategy. For details, see Materials and Methods.

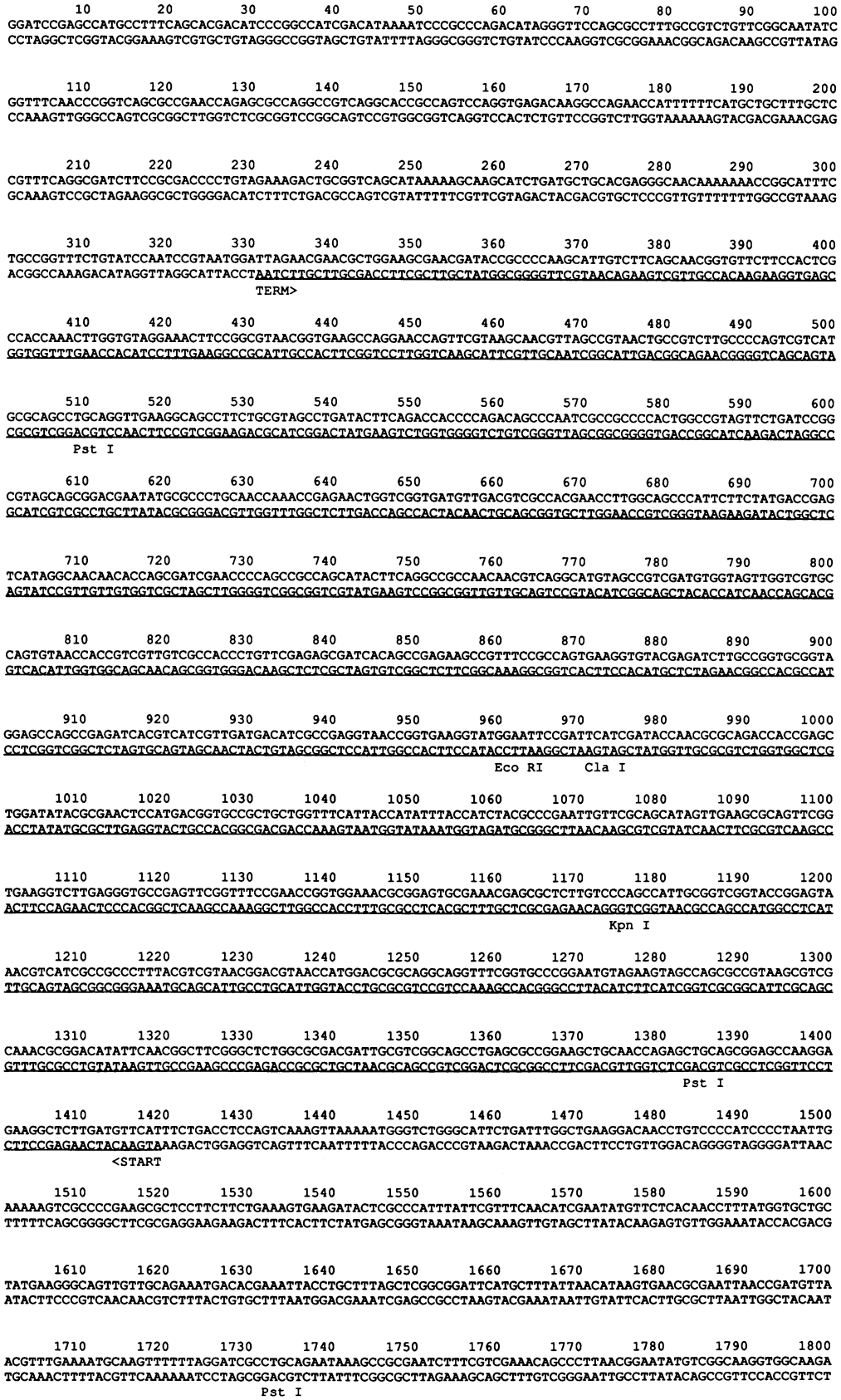
digested with several restriction enzymes. Single fragments of 6.5 kbp (S2308) and 3.5 kbp (S19) containing complete gene copies were identified in the *Bam*HI digests to which hybridized a λgt11 insert and a 5′-end-specific oligonucleotide probe (GAPuCCNGAPuGCNGT). These fragments were subcloned into pBR322 as described in Materials and Methods. The λgt11 probes were obtained previously by using antibody raised against the purified protein isolated from the cell envelopes following sodium dodecyl sulfate-polyacrylamide gel electrophoresis (7).

Duplicate Southern blots of the pAGF201 recombinant, digested with four restriction enzymes, were hybridized with the original λgt11 insert (Fig. 1A) or the oligonucleotide probe (Fig. 1B). The results indicate that the λgt11 probe hybridized to a single fragment in the *Bam*HI digest (Fig. 1, lanes 1, 7.5 kbp) and to two fragments each in the *Eco*RI digest (lanes 2, 4.8 and 1.3 kbp), *Pst*I digest (lanes 3, 890 and 650 bp), and *Cla*I digest (lanes 4, 2.3 and 1.3 kbp). The absence of *Eco*RI, *Pst*I, and *Cla*I restriction sites from the probed sequences indicated that a duplication of all or part of the 280 bp was present within a 3.5-kbp region (7).

The oligonucleotide probe hybridized to the same fragments which hybridized to the λgt11 probe in the *Bam*HI (Fig. 1, lanes 1, 7.5 kbp) and *Pst*I (lanes 3, 890 and 650 bp) digests, to only one of two fragments in the *Cla*I digest (lanes 4, 2.3 kbp), and to a different fragment altogether in the *Eco*RI digests (lanes 2, 1.6 kbp). Weak hybridization was observed with several other DNA fragments and was presumably a background level, as no other significant DNA homologies have been found following DNA sequence analysis. These results indicate the duplication of the sequence hybridizing to the oligonucleotide probe. In addition, hybridization to a single *Eco*RI fragment indicated that the repeated sequences were positioned fewer than 1,600 bp apart and presumably were flanked by the sequences hybridizing to the λgt11 probe. Since the oligonucleotide represented the 5′ end of the gene and the λgt11 sequences should be located downstream of these, the results suggest that the duplicated sequences were oriented in opposite directions. The results from the Smith and Birnstiel mapping (32) (Fig. 2) are consistent with the proposed gene organization, indicating that the *Eco*RI fragments of 1.3 and 4.8 kbp which hybridize to the λgt11 probe flank the 1.6-kbp *Eco*RI fragment.

The results from the *Pst*I digests (Fig. 1, lanes 3) indicate the presence of two similar but nonidentical regions, represented by one 890-bp fragment and one 650-bp fragment, which hybridized to both the oligonucleotide probe and the λgt11 probe. From the mapping data (Fig. 2) we saw that these two fragments were separated by two additional *Pst*I fragments of 350 and 550 bp; thus, the duplicated sequences were separated by no fewer than 900 bp but in the case of the λgt11 encoded portions, by as many as 1,800 bp. The latter distance is too great to represent a single gene encoding a 36-kDa protein containing repeated regions within 1,100 bp. Thus, the data are best explained by the presence of two closely related but nonidentical genes oriented in opposite directions and located within a 3.5-kbp stretch of genomic DNA. In order to clarify this arrangement, DNA sequence analysis was performed.

DNA sequence analysis. The DNA sequence of this region was determined (Fig. 2B) from overlapping fragments which were subcloned into M13 mp18 and mp19 and sequenced as described in Materials and Methods. The complete nucleotide sequence of this locus (3.5 kbp) was identical for both strains 19 and 2308 (Fig. 3). There were three large open reading frames which started with ATG initiation codons, only two of which had upstream Shine-Dalgarno sequences (GGAGG) and are considered in this section. The third large open reading frame was located on the top strand from positions 499 to 1425 and are discussed below. *omp* 2a was located on the top strand from positions 2251 to 3216, and *omp* 2b was located on the bottom strand from positions 1419 to 331 (Fig. 3, underlined). Sequences corresponding to the oligonucleotide probe (bp 1332 to 1319 and 2338 to 2351) and the λgt11 insert (bp 966 to 690 and 2532 to 2812) were identified in both regions and were oriented in different



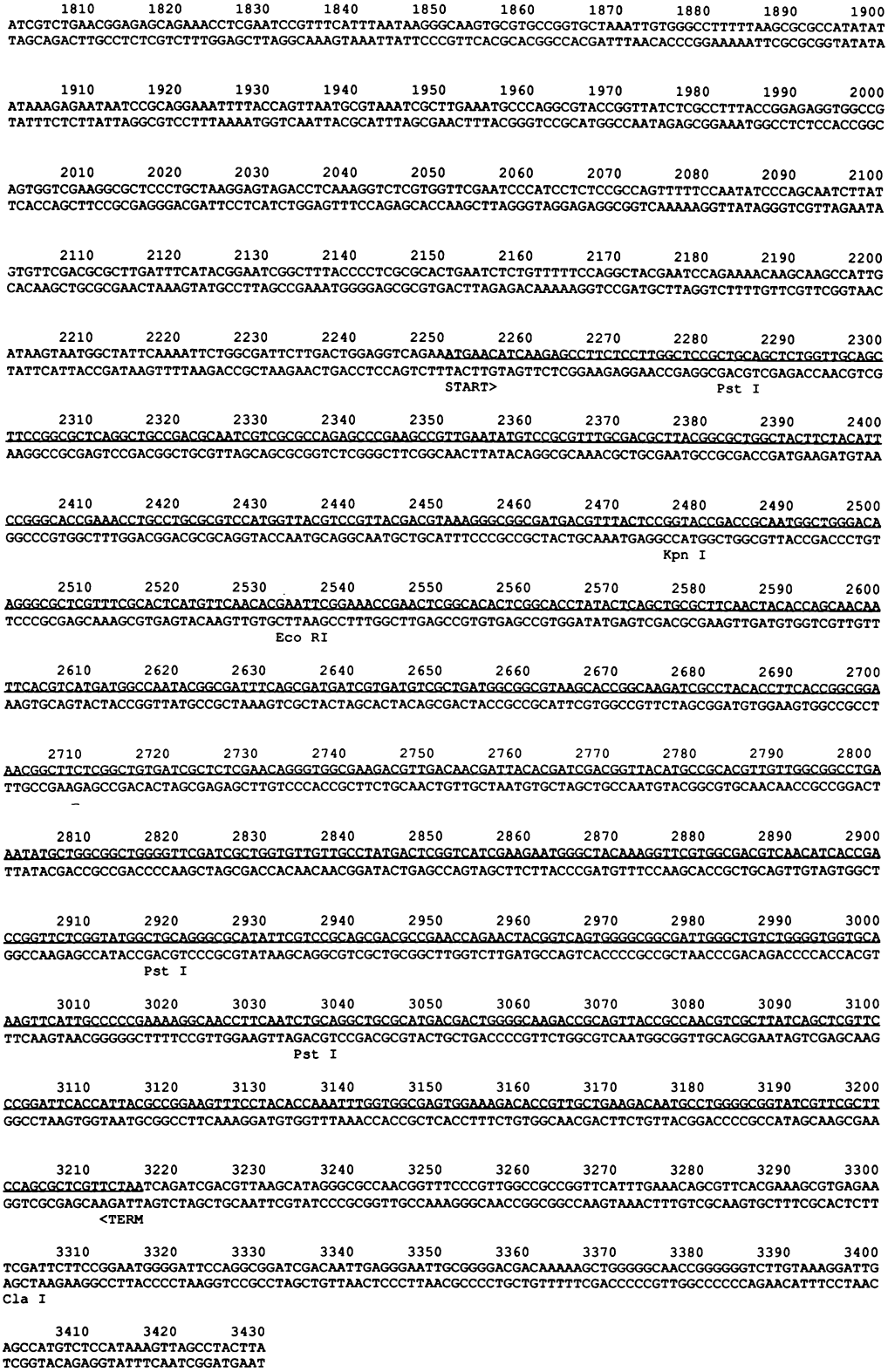


FIG. 3. DNA sequence of the *omp* 2 locus of *B. abortus* (GenBank no. M26034). The complete nucleotide sequence of the *omp* 2 locus was determined for S19 and S2308 as described in Materials and Methods. The *omp* 2a and *omp* 2b open reading frames are described in the text and are underlined.

directions. The homology between the λgt11 probe and *omp* 2b was perfect, whereas the homology with *omp* 2a was not. This presumably accounts for the difference in band intensity observed on Southern blots (Fig. 1A, lane 3).

Alignment of the two homologous regions, nucleotides 2184 to 3430 with 1288 to 117 (Fig. 4), revealed that the two regions shared 100% homology over the first 287 bp. The middle one-third of the two genes differed because of the absence of a 108-bp segment from *omp* 2a. Several smaller insertions and deletions (i.e., 3, 6, 9, and 18 bp) were distributed between both gene copies. In all cases except one, the insertions and deletions were multiples of three and, as a result, maintained the final reading frame. The single exception was the absence of 2+1 nucleotides over a stretch of 4 bp in *omp* 2b. This could easily be rewritten to indicate the removal of three contiguous bases; however, this would not conform with maximal sequence alignment. Whether the rearrangements have occurred in multiples of three in one or both gene copies is a moot point and must await analysis of the DNA of other *Brucella* species. The 3′ one-third of both genes exhibited 90% homology which extended only as far as the termination codon. Beyond the open reading frames, a potential rho-independent transcription termination site containing a GC-rich region capable of forming a stable hairpin upstream of a poly(U) stretch was found at the 3′ end of *omp* 2b (309 to 284) (Fig. 4, underlined). The 3′ end of *omp* 2a also had a potential rho-independent transcription termination site (3239 to 3277) which could form a GC-rich hairpin. The loop was followed by the sequence UUCAUUU. Although reminiscent of the typical poly(U) stretch characteristic of rho-independent transcription, this alteration may be the cause of differential expression of these genes that use the same promoter (25).

Identification of the canonical promoter sequences upstream from either gene proved more difficult (17). This was not an unexpected result, since porin genes, such as *ompF* and *ompC* in *E. coli*, are positively regulated and, as such, have poor consensus promoter sequences (21, 26). Previous work has already demonstrated the presence of canonical promoter sequences upstream from some *Brucella* genes (16). As a result, characterization of promoter activity was performed via a functional assay, as described in the next section.

The predicted gene products, designated Omp 2a and Omp 2b, were 96% homologous over their shared length (Fig. 5). However, Omp 2a was shorter by 41 amino acids because of the missing nucleotides in the middle of the gene, and this resulted in an overall difference in their molecular masses, i.e., 38,720 Da for Omp 2b and 34,532 Da for Omp 2a. The mature end of the 36-kDa protein has been determined in one of our labs and had the sequence Ala-Asp-Ala, etc. (B. Sowa, unpublished results). Thus, both genes appeared to encode proteins with 22-amino-acid-long signal peptides which were cleaved between alanine residues in the sequence Ala-Ala-Asp-Ala, reminiscent of *E. coli* signal peptide cleavage sites (22).

The original antisera used to select these recombinants were raised against a major protein identified in cell envelopes isolated from *B. abortus* which we believe represents the outer membrane porin. Comparison of the amino acid composition of either gene product predicted from the DNA sequence with that determined for the *Brucella* porin isolated by Verstreate and co-workers (35) revealed almost perfect identity (Table 1). We believe that these data are consistent with the identification of the *omp* 2 locus of *B. abortus*, which contained a unique arrangement of porin genes which we tentatively designated *omp* 2a and *omp* 2b (Fig. 4 and 5).

Transcriptional activity at the *omp* 2 locus. Since a single protein species reactive with the antisera was detected in the cell envelopes of laboratory-grown *B. abortus*, it was important to determine if both porin genes were functional. S1 mapping was not used to pinpoint the transcription start sites, since the homology between the two gene transcripts would not allow a distinction between one or two active promoters. The upstream *Pst*I fragments of 350 bp (*omp* 2b) and 550 bp (*omp* 2a) were cloned upstream of and fused in frame to a promoterless *lacZ* gene in pMC1871 (4). These plasmids were transformed into *E. coli* MC4100, and the β-galactosidase activity was measured as described by Miller (20). β-Galactosidase activity was detected only in cells carrying the 350-bp fragment in an orientation consistent with its arrangement in *omp* 2b (Fig. 6). Verification of insert orientation was performed by restriction enzyme analysis as described in Materials and Methods. No activity was detected above background level in cells containing a plasmid carrying the 550-bp fragment in either orientation. These data suggest the presence of a single functional promoter capable of expressing activity in *E. coli*. This does not rule out possible expression from the 550-bp fragment under other growth conditions or only in *B. abortus*. Experiments designed to test the environmental regulation of expression from the *omp* 2 locus on the basis of changes in β-galactosidase levels under variant conditions of osmolarity and temperature had no effect on *E. coli* (data not shown). However, the effect of these factors on expression from this locus in *B. abortus* has not been examined and cannot be ruled out.

Protein expression from the *omp* 2 locus. In order to determine whether there are additional blockages to the expression of *omp* 2a, the orientation of the unique *Kpn*I fragment (Fig. 3, positions 1189 to 2476) was inverted relative to the native configuration, thus placing *omp* 2a expression under the control of the active *omp* 2b promoter. The cell envelope fraction from *E. coli* ECB611 containing plasmids with the *Kpn*I fragment in the native (pAGF101 and pAGF201) or inverted (pAGF111 and pAGF211) configuration or deleted (pAGF11 and pAGF1121) were examined by Western blot analysis. These cells were used because of the increased expression of *omp* 2 gene products compared with that produced in the porin-replete parental strain, MC4100. When the *Kpn*I fragment was in the native configuration, a protein of approximately 36 kDa was detected (Fig. 7A, lanes 1 and 4), and when the *Kpn*I fragment was inverted, a slightly smaller protein of 33 kDa was observed (Fig. 7A, lanes 2 and 5). In the absence of the *Kpn*I fragment, production of either protein was undetectable (Fig. 7A, lanes 3 and 6). Since expression in either the native or inverted configuration is controlled by a single promoter, the difference in the level of expression may reflect differences in mRNA stability or protein stability or differences in antibody binding to the proteins which, although related, do have regions which are unique. In order to rule out the relative antibody affinities for these two proteins, synthesis in maxicells was examined. The results indicate that the level of Omp 2a (Fig. 8, lanes 3 and 6) found in these cells was much lower than that for Omp 2b (Fig. 8, lanes 2 and 5). The reasons for this are not known but may be related to differences in RNA production or stability or in protein stability (see Discussion). Some of this difference is attributable to the difference in amino acid content of the two proteins, since Omp 2b would have a higher specific activity

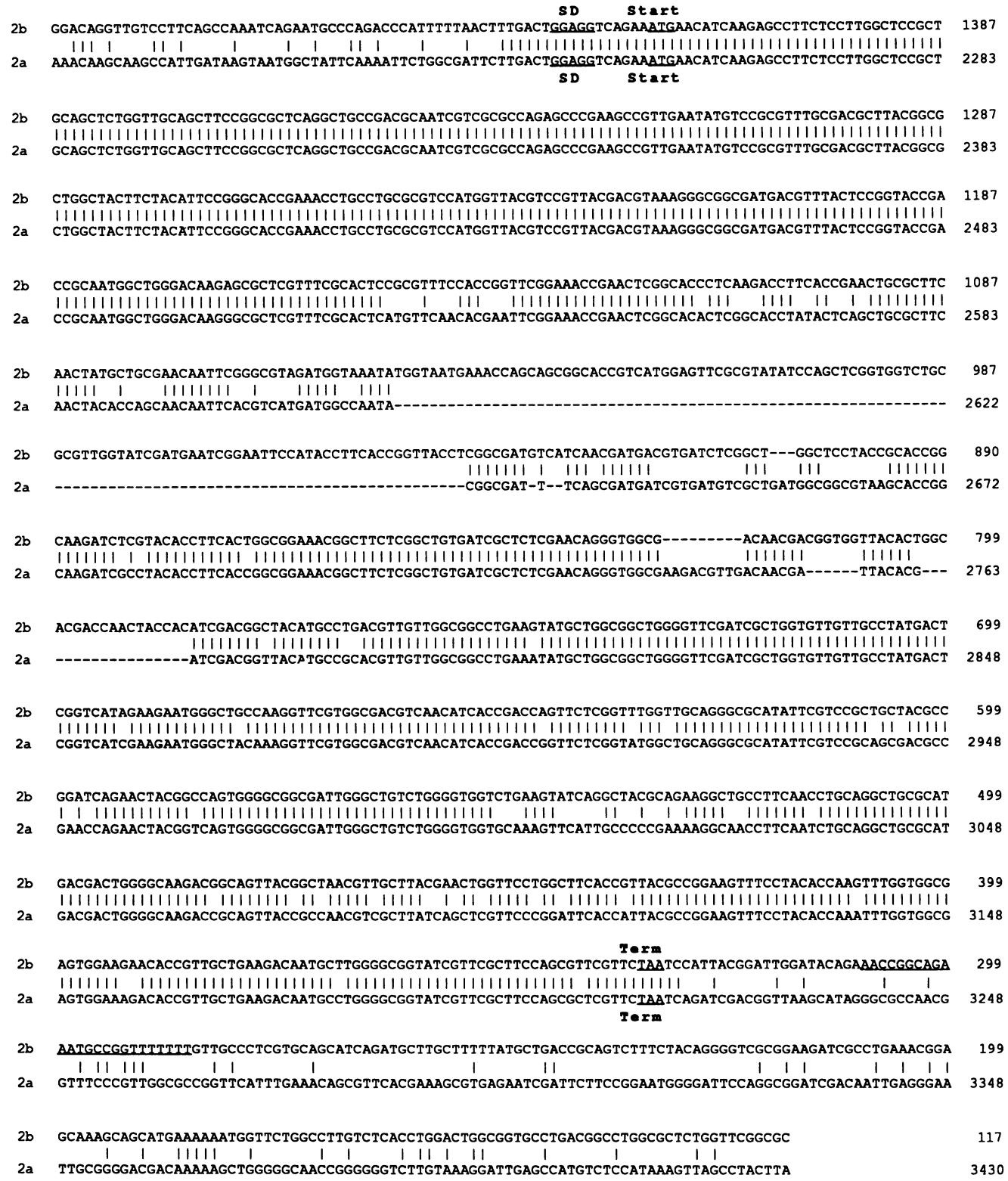


FIG. 4. Alignment of *omp* 2a and *omp* 2b. The open reading frames corresponding to *omp* 2a and *omp* 2b are shown in alignment following inversion of the *omp* 2b sequences as described in the text. |, Homology (nonhomology is indicated by the absence of a mark). SD, Shine-Dalgarno sequence.

```
Omp 2b  MNIKSLLLGSAAALVAASGAQAADAIVAPEPEAVEYVRVCDAYGAGYFYIPGTETCLRVH    60aa
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Omp 2a  MNIKSLLLGSAAALVAASGAQAADAIVAPEPEAVEYVRVCDAYGAGYFYIPGTETCLRVH    60aa
                                  ◊

Omp 2b  GYVRYDVKGGDDVYSGTDRNGWDKSARFALRVSTGSETELGTLKTFTELRFNYAANNSGV    120aa
        ||||||||||||||||||||||||*|||||***|*|||||||||*|*|*|||||**|||**
Omp 2a  GYVRYDVKGGDDVYSGTDRNGWDKGARFALMFNTNSETELGTLGTYTQLRFNYTSNNSRH    120aa

Omp 2b  DGKYGNETSSGTVMEFAYIQLGGLRVGIDESEFHTFTGYLGDVINDDVISAGSYRTGKIS    180aa
        ||*||***                                    *|*****|***||||*
Omp 2a  DGQYGDFS------------------------------------DDRDVADGGVSTGKIA    144aa

Omp 2b  YTFTGGNGFSAVIALEQGGDNDGGYTGTTNYHIDGYMPDVVGGLKYAGGWGSIAGVVAYD    240aa
        ||||||||||||||||||******     |*||||||*||||||||||||||||||||||
Omp 2a  YTFTGGNGFSAVIALEQGGEDVDND-----YTIDGYMPHVVGGLKYAGGWGSIAGVVAYD    199aa

Omp 2b  SVIEEWAAKVRGDVNITDQFSVWLQGAYSSAATPDQNYGQWGGDWAVWGGLKYQATQKAA    300aa
        |||||||*|||||||||*|||||||||||||||*|||||||||||||||*|**|**||*
Omp 2a  SVIEEWATKVRGDVNITDRFSVWLQGAYSSAATPNQNYGQWGGDWAVWGGAKFIAPEKAT    259aa

Omp 2b  FNLQAAHDDWGKTAVTANVAYELVPGFTVTPEVSYTKFGGEWKNTVAEDNAWGGIVRFQRSF    362aa
        ||||||||||||||||||||*||||||*|||||||||||||||*|||||||||||||||||
Omp 2a  FNLQAAHDDWGKTAVTANVAYQLVPGFTITPEVSYTKFGGEWKDTVAEDNAWGGIVRFQRSF    321aa
```

FIG. 5. Alignment of the gene products Omp 2a and Omp 2b. The putative products of *omp* 2a and *omp* 2b are shown in alignment. |, Homologous position; *, nonhomology; ◇, putative signal peptide cleavage point.

than Omp 2a (Omp 2a, Gly + Leu = 69; Omp 2b, Gly + Leu = 57). However, this cannot account for the large difference observed.

The only block to expression from the *omp* 2a in *E. coli* appeared to be the absence of a functional promoter. However, even when linked to an active promoter, *omp* 2a was poorly expressed. Whether this is also the case in *B. abortus* remains to be demonstrated. However, it is clear that only a single species of porin was expressed under normal conditions in laboratory-grown *Brucella* spp. (Fig. 7B). This protein comigrated with *omp* 2b gene product produced in *E. coli* and was consistent with the demonstrated promoter activity of this gene.

Additional protein bands which appeared to be synthesized from the cloned *Brucella* DNA include protein 2d, which originated from the region unique to the larger insert in pAGF101 (Fig. 8, lanes 1 through 3) and was missing from pAGF201 (Fig. 8, lanes 4 through 6) recombinants. This region has not been sequenced, and we do not know the nature or the exact origin of this gene product. The protein labeled 2c was the size predicted for the product of the third large open reading frame, positions 499 to 1425. However, the protein appeared to be synthesized, although at reduced levels, in *E. coli* containing pAGF11 and pAGF21, which have deletions spanning nucleotides 1189 to 2476 which would fuse this protein to the *omp* 2a gene product. Thus,

TABLE 1. Comparison of the predicted amino acid compositions of *omp2* gene products with that determined for the *B. abortus* porin[a]

Amino acid	No. (%) occurring in:		% Occurring in *B. abortus* group 2 porin[b]
	Omp 2a	Omp 2b	
Acid			
Ala	37 (11.5)	40 (11.0)	10.8
Gly	43 (13.4)	51 (14.1)	15.4
Val	28 (8.7)	32 (8.8)	8.1
Leu	14 (4.4)	18 (5.0)	5.3
Ile	12 (3.7)	14 (3.9)	4.0
Pro	8 (2.5)	7 (1.9)	2.4
Phe	14 (4.4)	15 (4.1)	4.5
Trp	10 (3.1)	10 (2.8)	ND
Met	3 (0.9)	3 (0.8)	0.6
Ser	17 (5.3)	22 (6.1)	5.9
Thr	23 (7.2)	27 (7.5)	7.7
Cys	2 (0.6)	2 (0.6)	ND
Tyr	19 (5.9)	23 (6.4)	6.1
Asn	15 (4.7)	16 (4.4)	—
Gln	10 (3.1)	11 (3.0)	—
Lys	11 (3.4)	13 (3.6)	4.0
Arg	12 (3.7)	12 (3.3)	3.8
His	4 (1.2)	4 (1.1)	1.2
Asp	25 (7.8)	24 (6.6)	—
Glu	14 (4.4)	18 (5.0)	—
Asx	40 (12.5)	40 (11.0)	11.7
Glx	24 (7.5)	29 (8.0)	9.4
Type			
Acidic	39 (12)	42 (12)	
Basic	27 (8)	29 (8)	

[a] Molecular masses for Omp 2a, Omp 2b, and *B. abortus* group 2 porin were 34.5, 38.7, and 35 to 41 kDa, respectively.

[b] Data from reference 35. ND, Not determined; —, values for Asn and Asp for Gln and Glu were not separately determined.

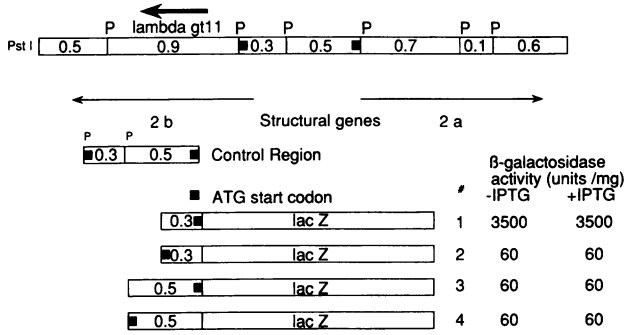


FIG. 6. Identification of active promoters by fusion to a promoterless *lacZ* gene. Construction of recombinant plasmids fusing the putative *omp* 2a and *omp* 2b promoters to *lacZ* was performed as described in the text. β-Galactosidase activity produced by these plasmids was measured as described in Materials and Methods. IPTG, Isopropyl-β-D-thiogalactopyranoside.

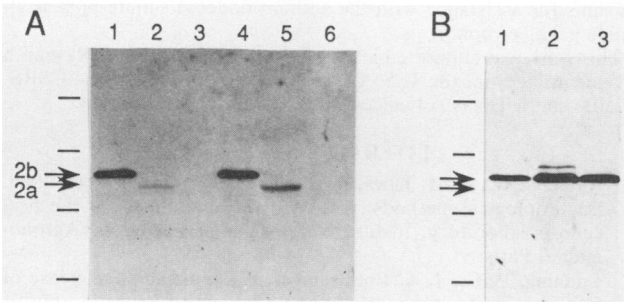


FIG. 7. Western blot analysis of *omp* 2a and *omp* 2b gene products synthesized in *E. coli*. *Brucella* proteins were synthesized in *E. coli* ECB611, a porin-deficient derivative of MC4100. The cell envelope fraction was collected, and the presence of the *omp* 2 gene products was determined by Western blot as described in Materials and Methods (A) pAGF101 (lane 1), pAGF111 (lane 2), pAGF11 (lane 3), pAGF201 (lane 4), pAGF211 (lane 5), and pAGF21 (lane 6); (B) pAGF201 (lane 1), *B. abortus* S2308 (lane 2), and *B. abortus* S19 (lane 3). Horizontal lines indicate the migration of molecular mass markers of 66 kDa (bovine serum albumin), 43 kDa (ovalbumin), 31 kDa (carbonic anhydrase), and 21 kDa (soybean trypsin inhibitor), from top to bottom. Omp 2a and 2b are indicated by arrows. Plasmids are described in the text.

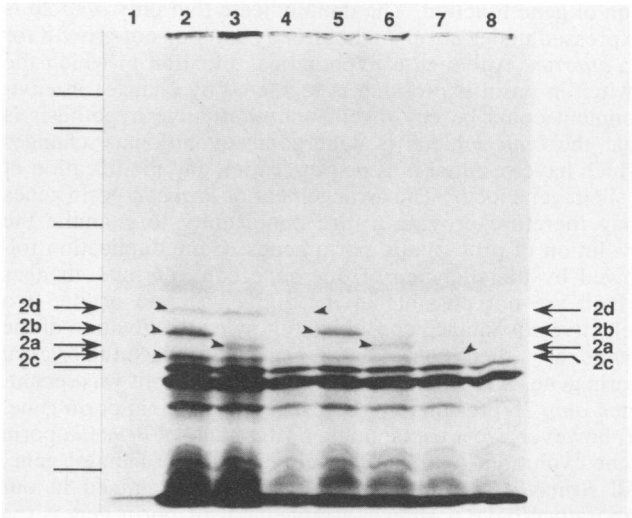


FIG. 8. *B. abortus omp* 2-directed protein synthesis in maxicells. Maxicells were prepared and proteins were electrophoresed as described in the text. Lanes: 1, *E. coli* SE5000; 2, pAGF101; 3, pAGF111; 4, pAGF11; 5, pAGF201; 6, pAGF211; 7, pAGF21; 8, pBR322. Horizontal lines indicate the migration of molecular mass markers of 116 kDa (β-galactosidase), 97.4 kDa (phosphorylase *b*), 66 kDa (bovine serum albumin), 43 kDa (ovalbumin), 31 kDa (carbonic anhydrase), and 21 kDa (soybean trypsin inhibitor), from top to bottom. Plasmids are described in the text.

the origin of protein 2c has not been established at this time and may simply represent an artifactual expression of sequences caused by fusion of *Brucella* and pBR322 DNAs.

DISCUSSION

The *omp* 2 locus of *B. abortus* was shown to be composed of two homologous regions. One contained the *omp* 2b gene, which was active in *E. coli* and encoded a protein of 36 kDa which comigrated with a major outer membrane protein identified in all laboratory-grown *Brucella* spp. (7). The predicted amino acid composition was virtually identical to that determined for the *Brucella* group 2 outer membrane proteins which demonstrate porin activity (6, 35). Secondary-structure predictions based on the rules of Chou and Fasman (5) and Garnier et al. (9) indicated that the Omp 2b protein has extensive regions of β-sheet and β-turn (data not shown). In addition, the predicted protein exhibited alternating hydrophilic and hydrophobic domains necessary for pore formation across a lipid bilayer. Although by no means definitive, all of the evidence is consistent with the identification of the *Brucella* porin gene.

The second region contained an open reading frame temporarily identified as the *omp* 2a gene, which was homologous to *omp* 2b but was inactive in *E. coli*. When linked to an active promoter, the gene encoded a 33-kDa protein with amino acid composition and structural features similar to those of the *omp* 2b gene product. However, no evidence of such a protein could be detected in laboratory-grown *Brucella* spp. Thus, the *omp* 2a gene, by definition, represents a silent copy of *omp* 2b. This raises questions concerning the origin and potential role of both genes and their encoded products in the survival and virulence of *B. abortus*.

Regulated divergent transcription. The most likely explanation for this gene arrangement is that the *omp* 2a and *omp* 2b genes represent the *Brucella* counterparts of *E. coli ompF* and *ompC* and that the expression of *omp* 2a is regulated by some environmental condition unspecified as yet (26). Experiments in progress are designed to examine the outer membrane composition of *Brucella* spp. isolated from infected cattle or following growth under different conditions in vitro. In addition, the abilities of both gene products to complement porin-deficient *E. coli* are currently being investigated. As demonstrated for *E. coli* and several other procaryotes, a highly conserved system is used to sense changes in environmental conditions and to induce a switch in gene expression at the transcriptional level. This system has been modified and used by a number of different bacteria in order to detect and respond to environmental conditions (28). For the brucellae, the nature of those environmental conditions is not known but is suspected to be related to the ability of this organism to survive intracellularly in professional phagocytes.

The arrangement of *omp* 2a and *omp* 2b is conducive to regulation which could potentially express one gene or the other simply by shifting the direction of transcription. The mechanism by which this is achieved can only be postulated but need not differ much from that which is proposed to occur in *E. coli* in the regulation of *ompF* and *ompC* expression, as described above (26, 31). One difference is that the *E. coli* porin genes are separated by over 10^6 bp and would require forms of regulation different from those that could be used to regulate porin gene expression in *Brucella* spp. For example, regulation of *omp* 2a and *omp* 2b expression could utilize a switching mechanism to invert the upstream sequence, bringing the *omp* 2b promoter into proper alignment with the *omp* 2a gene as was performed artificially for our experiment. Examples of this kind of mechanism, such as the phase variation in *Salmonella* spp. and in *E. coli*, abound in the procaryotes (24). However, consensus sequences described for these mechanisms were not found in the region upstream of either structural gene.

Evolution of *B. abortus* porin genes. Evidence for the occurrence of DNA rearrangements within the *omp* 2 structural genes following duplication has been presented. Despite extensive changes in sequence, the conservation of reading frame in both genes suggests the potential preserva-

tion of gene function. The data indicate that only *omp* 2b is expressed under normal laboratory conditions of growth for *B. abortus*. Although a hypothetical situation in which the switch in porin expression is regulated by changes in environment could be envisaged, one alternative hypothesis is that the *omp* 2 locus is undergoing evolutionary changes which have resulted in gene duplication and modification of a silent gene locus. The arrangement of *Brucella* porin genes may therefore provide a rare opportunity to examine the evolution of procaryotic porin genes. Gene duplication followed by the silencing of one gene copy permits changes which are not possible in the active gene to occur. If a selective advantage is provided following reactivation of the silent copy, then the new porin gene will replace the original porin gene. Assuming that the duplication event was recent, then *omp* 2b presumably represents the original porin gene. If, however, we are examining a later stage of *Brucella* porin gene evolution, *omp* 2a may represent the progenitor gene. All *Brucella* species and strains so far examined in our laboratory have a similar arrangement of porin genes: the exact nature of these remains to be determined (T. A. Ficht, unpublished results). This is consistent with the hypothesis that gene duplication occurred prior to speciation. Speciation within the genus *Brucella* is not clearly defined, and all species share greater than 95% overall DNA homology, consistent with a monospecific genus (34). Thus, it cannot be stated with certainty whether *omp* 2a represents the inactive remnant of a once-functional porin gene, randomly undergoing change which may lead to improved porin function or eventual deletion from the genome, or a highly conserved gene which is active only under specified conditions.

A third possibility is that *omp* 2a represents a cryptic gene, the function of which is necessary under certain growth conditions but is detrimental to the growth of the cell under other growth conditions. One example of a cryptic gene is the *bgl* operon of *E. coli*, which permits the utilization of β-glucosides as the sole carbon and energy source (23). However, cryptic genes are not normally duplicated. Instead, mechanisms have evolved to silence these genes under normal conditions, and only under unusual conditions do mutations which activate expression provide a selective advantage (10). The evolution of cryptic gene function can proceed without duplication, since cryptic genes are rarely expressed. According to this definition, the *omp* 2a gene is not expected to fall into the category of a cryptic gene.

The *omp* 2 genes described have a unique arrangement unlike any previously described for porins of gram-negative bacteria. The closest examples are the protein P porin genes of *Pseudomonas aeruginosa*, which are arranged in a tandem duplication (30). In this case both copies of the gene appear identical; however, only one gene copy is expressed. The significance of the head-to-head arrangement of *Brucella* porin genes must await further characterization. Although experiments described here indicate the absence of expression from *omp* 2a in *E. coli*, this cannot be extended a priori to *Brucella* spp., which may have evolved a positive regulatory mechanism for expression from *omp* 2a. Although suggestive, the data presented above cannot resolve the conflicting hypotheses of regulated divergent transcription versus silent gene.

ACKNOWLEDGMENTS

We thank Doug Struck, Roger Smith III, and Allison Ficht for their critical reading of the manuscript. We also thank Debby Jones and Christine R. Wendroch for technical assistance and Hélène Marquis for assistance with the sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

This work was supported in part by a Cooperative State Research service grant from the U.S. Department of Agriculture (84-CSRS-2-2503) and a Texas Advanced Technology Research grant.

LITERATURE CITED

1. **Alton, G. G., L. M. Jones, R. D. Angus, and J. M. Verger.** 1988. Bacteriological methods, p. 13–61. *In* Techniques for the brucellosis laboratory. Institut National de la Recherche Agronomique, Paris.
2. **Canning, P. C., J. A. Roth, and B. L. Deyoe.** 1986. Release of 5′-guanosine monophosphate and adenine by *Brucella abortus* and their role in the intracellular survival of the bacteria. J. Infect. Dis. **154**:464–470.
3. **Canning, P. C., J. A. Roth, L. B. Tabatabai, and B. L. Deyoe.** 1985. Isolation of components of *Brucella abortus* responsible for inhibition of function in bovine neutrophils. J. Infect. Dis. **152**:913–921.
4. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143**:971–980.
5. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. **47**:45–148.
6. **Douglas, J. T., E. Y. Rosenberg, H. Nikaido, D. R. Verstreate, and A. J. Winter.** 1984. Porins of *Brucella* species. Infect. Immun. **44**:16–21.
7. **Ficht, T. A., S. W. Bearden, B. A. Sowa, and L. G. Adams.** 1988. A 36-kilodalton *Brucella abortus* cell envelope protein is encoded by repeated sequences closely linked in the genomic DNA. Infect. Immun. **56**:2036–2046.
8. **Frost, A. J., H. W. K. Smith, and J. Keppie.** 1972. The chemical basis of the virulence of *Brucella abortus* X. A surface virulence factor which facilitates intracellular growth of *Brucella abortus* in bovine phagocytes. Br. J. Exp. Pathol. **53**:587–596.
9. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**:97–120.
10. **Hall, B. G., S. Yokoyama, and D. H. Calhoun.** 1983. Role of cryptic genes in microbial evolution. Mol. Biol. Evol. **1**:109–124.
11. **Hu, N., and J. Messing.** 1982. The making of strand-specific M13 probes. Gene **17**:271–277.
12. **Kreutzer, D. L., and D. C. Robertson.** 1979. Surface macromolecules and virulence in intracellular parasitism: comparison of cell envelope components of smooth and rough strains of *Brucella abortus*. Infect. Immun. **23**:819–828.
13. **Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
14. **Lutkenhaus, J. F.** 1977. Role of a major outer membrane protein in *Escherichia coli*. J. Bacteriol. **131**:631–637.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 545. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Mayfield, J., B. Bricker, H. Godfrey, R. Crosby, D. Knight, S. Halling, D. Balinsky, and L. Tabatabai.** 1988. The cloning, expression, and nucleotide sequence of a gene coding for an immunogenic *Brucella abortus* protein. Gene **63**:1–9.
17. **McClure, W.** 1985. Mechanism and control of transcription initiation in prokaryotes. Annu. Rev. Biochem. **54**:171–204.
18. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19**:269–276.
19. **Meyer, M. E.** 1967. Metabolic characterization of the genus *Brucella*. VI. Growth stimulation by *i*-erythritol compared with strain virulence for guinea pigs. J. Bacteriol. **93**:996–1000.
20. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

21. **Mizuno, T., M.-Y. Chou, and M. Inouye.** 1983. A comparative study on the genes for three porins of the *E. coli* outer membrane: DNA sequence of the osmoregulated *omp* C gene. J. Biol. Chem. **258**:6932–6940.
22. **Mutoh, N., K. Inokuchi, and S. Mizushima.** 1982. Amino acid sequence of the signal peptide of Omp F, a major outer membrane protein of *Escherichia coli*. FEBS Lett. **137**:171–174.
23. **Parker, L. L., and B. G. Hall.** 1988. A fourth *Escherichia coli* gene system with the potential to evolve β-glucoside utilization. Genetics **119**:485–490.
24. **Plasterk, R. H. A., A. Brinkman, and P. van de Putte.** 1984. DNA inversions in the chromosome of *Escherichia coli* and in bacteriophage Mu: relationship to other site-specific recombination systems. Proc. Natl. Acad. Sci. USA **80**:5355–5358.
25. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55**:339–372.
26. **Ramakrishnan, G., D. E. Comeau, K. Ikenaka, and M. Inouye.** 1987. Transcriptional control of gene expression: osmoregulation of porin protein synthesis, p. 3–16. *In* M. Inouye (ed.), Bacterial outer membranes as model systems. John Wiley & Sons, Inc., New York.
27. **Riley, L. K., and D. C. Robertson.** 1984. Ingestion and intracellular survival of *Brucella abortus* in human and bovine polymorphonuclear leukocytes. Infect. Immun. **46**:224–230.
28. **Ronson, C. W., B. T. Nixon, and F. M. Ausubel.** 1987. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. Cell **49**:579–581.
29. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. **137**: 692–693.
30. **Siehnel, R. J., E. A. Worobec, and R. E. W. Hancock.** 1988. Cloning of the *Pseudomonas aeruginosa* outer membrane porin protein P gene. J. Bacteriol. **170**:2312–2318.
31. **Slauch, J. M., S. Garrett, D. E. Jackson, and T. J. Silhavy.** 1988. EnvZ functions through OmpR to control porin gene expression in *Escherichia coli* K-12. J. Bacteriol. **170**:439–441.
32. **Smith, H. O., and M. L. Birnstiel.** 1976. A simple method for DNA restriction site mapping. Nucleic Acids Res. **3**:2387–2398.
33. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
34. **Verger, J. M., F. Grimont, P. A. D. Grimont, and M. Grayon.** 1985. *Brucella*, a monospecific genus as shown by deoxyribonucleic acid hybridization. Int. J. Syst. Bacteriol. **35**:292–295.
35. **Verstreate, D. R., M. T. Creasy, N. T. Caveney, C. L. Baldwin, M. W. Blab, and A. J. Winter.** 1982. Outer membrane proteins of *Brucella abortus*: isolation and characterization. Infect. Immun. **35**:979–989.
36. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, and M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.